1/10

entry clone

4088 bp

T1-T2

ori

CDS0689

Selectable marker

FIGURE 1

CDS0689

T-zein
T-rbcS-deltaGA

TATA Signal 1

CAAT 1

prolamine

tNOS

RB repeat nopaline

RB Ti C58

prm0131

prim0130

screenable marker

constitutive promoter
- screenable marker
tnos cassette

Expression vector

12836 bp

constitutive promoter

polyA

pBR322 (ori + bom)

selectable marker 3' exon

constitutive promoter -
selectable marker (intron) -
polyA cassette

Sm/SpR

selectable marker 5' exon

TATA box

LB Ti C58

LB repeat nopaline

constitutive promoter

FIGURE 2

2/10

```
                                             1            Motif 1              40
                    CDS0689   (1)  ----MSVLQYPEGIDPADVQIWNNAAFDNGDSEDLSBLKR
                 CDS0689 At   (1)  ---MTSIEATETLNAPPKIQIWNNAAFDDGDSQIISAIEA
 CDS0689 Medicago trunculata  (1)  ---MNNTNNNNILLHSTQVQVINNAAFGEDFAMNSSSDS
              CDS0689 Os      (1)  ---------MEEDPLIPLVHVWNNAAFDDSSCSRSAWIPQ
           CDS0689 Ta variant (1)  --------MMEEDPLIPLVHVNTNAAFSSSSSSSAWHAHA
              CDS0689 So      (1)  --------MEEDPLIPLVHVINNAAFDHASSSAWHAHSP
        CDS0689 Zm variantrev (1)  --------MEEDPLIQLVHVISNACDNAARSSSVCHAH
           CDS0689 So variant (1)  --------MEEDPLIPLVHVWNNAPFHASYSAWHAHSP
            CDS0689 Bn est1    (1)  ----MTSTEHTETLNAPELQIWNNAAFDDGDSNLISAIEA
CDS0689 Eschscholzia californica (1) ------MLEISETLNIPDIQIWNNAAFSGSIDNHTTAIK
            CDS0689 Ga est1    (1)  ----MSILQYPDSFNVPELQVWNNAAFNGDSEDINAIKD
              CDS0689 Pt       (1)  ---MSSILQYPDVVDAPEVQIWNNAAFNGESEGSLNIKS
    CDS0689 Plumbao zeylanica  (1)  MNEVLHIQEAARTDSSTDHQIWNNAAFSGESEDSPVIID
      CDS0689 Citrus sinensis  (1)  ----MSVLQYPITLNGQELQIWNNAAFINGESEDSTAMKG
                  Consensus    (1)        I     E    IPEVQIWNNAAFD GDS S AI

                                             41            Motif 2              80
                    CDS0689  (37)  --SWSPLKPLSVRPSDSFESDLSSKENQTILFENSSVNLS
                 CDS0689 At  (38)  SSWSHLN-------ESFDSDCS--KENQFIISVSSSLQSS
 CDS0689 Medicago trunculata (38)  --------------------IEELNFSAFN------
              CDS0689 Os     (32)  ---SP----------AVAAVRKGDKENHRIEVVD------
           CDS0689 Ta variant(33)  -----------------TPVRRGEKIRRLAETN------
              CDS0689 So     (32)  -----------VPASIRRLAEGDKLNHRLDPDP------
        CDS0689 Zm variantrev(32)  -----------SPAPASARIGIGDKNLRLEPDV------
           CDS0689 So variant(32)  -----------ARASAGHLAEGDKINHRLDPDP------
            CDS0689 Bn est1  (37)  ----------SWSNLNASFDSDCSIKEQILVSVSSSLKSS
CDS0689 Eschscholzia californica(35) ASSSPLKPIVLNQSEPSILDSIYTIEQTISCCISPVRTK
            CDS0689 Ga est1  (37)  SWCNFNS-----GSVNQSLESDGSKIQSILWIKSPVSFK
              CDS0689 Pt     (38)  -----------SWWNQSLESDASKINLSIVCEQSSPVFV
    CDS0689 Plumbao zeylanica(41)  --F------SAPNLSQSLLSDSSIKLNLSKSLAEMPHPAK
      CDS0689 Citrus sinensis(37)  SWANLKS-----VYMNQSLESDCSKILSERLNKSPTSSL
                  Consensus  (41)              S  ESDG KEN  P      S

                                             81                                120
                    CDS0689  (75)  SPLPIKPLNPNGALENSRLKPNKPNSKQSLDEMAARKSGK
                 CDS0689 At  (69)  VSITEAPS--------------AKSKTVKTKSAADRSKK-
 CDS0689 Medicago trunculata(49)  ---IVPS------------------------------SNK-
              CDS0689 Os     (53)  ---VAAG---------------------------------
           CDS0689 Ta variant(50)  ----------------------------------------
              CDS0689 So     (54)  ----------------------------------------
        CDS0689 Zm variantrev(55)  ----------------------------------------
           CDS0689 So variant(54)  ----------------------------------------
            CDS0689 Bn est1  (67)  VSFSTDDP---------------IRCGKVK-----EKPH
CDS0689 Eschscholzia californica(75) SPLPIKPLHPNG----------------------------
            CDS0689 Ga est1  (72)  STASVVKP-------LSSKNVTGNTREPFSAKMKSGVCKE
              CDS0689 Pt     (66)  NSSKPAKPLQ------------------------------
    CDS0689 Plumbao zeylanica(73)  SPMQK-----------------------------------
      CDS0689 Citrus sinensis(72)  KSCVPNKPLQVN---------SSVKNSQMKQLK--SVSK
                  Consensus  (81)  

                                            121       Motif 3: coiled coil    160
                    CDS0689 (115)  GNDFRDEKKIDEEIEEIQMEISRLSSRLEALRIEKAEKTV
                 CDS0689 At  (94)  -------RDIDAEIEEVEKEIGRLSIKLESLRIEKAEQTA
 CDS0689 Medicago trunculata(56)  -------RTIDDEIAEIESEIKRLISKLELLRVEKAERKI
              CDS0689 Os     (57)  -------YDVEAEIGHIEAEILRLSSRLHHLRVSKQPEPN
           CDS0689 Ta variant(50)  --------DADAEIARIEAEILRLSSRLHHLRVSKGHDAK
              CDS0689 So     (54)  --------DVEAEIGHIEAEILRLXSRLHHLRTSKQSEPS
        CDS0689 Zm variantrev(55)  ----------EEMRHIEAEILRLSLRLHHLRTSQQLQPP
           CDS0689 So variant(54)  --------DVEAEIGHIEAEILRLSSRLHHLRTSKQSEPP
            CDS0689 Bn est1  (86)  KTGKVRHGDIDAEIEEVEKFMNRLSIRLESLRIEKAEQIA
CDS0689 Eschscholzia californica(87) ----------------------------------------
            CDS0689 Ga est1 (105)  EEKKRDEKKIDMEIEEIEKEVARLSAKLESLRIEKPNIMQ
              CDS0689 Pt     (76)  ----------------------------------------
    CDS0689 Plumbao zeylanica(78)  ----------------------------------------
      CDS0689 Citrus sinensis(100) EEETRDERKIDIEIEEIEKEISRLSSRLEALRIEKIDIKT
                  Consensus (121)        ID EI  IE EI RLSSRL  LRL K
```
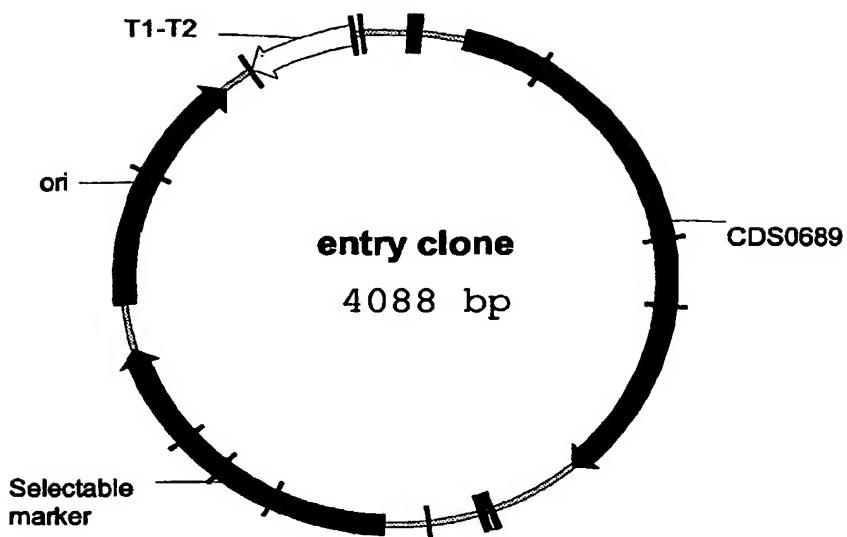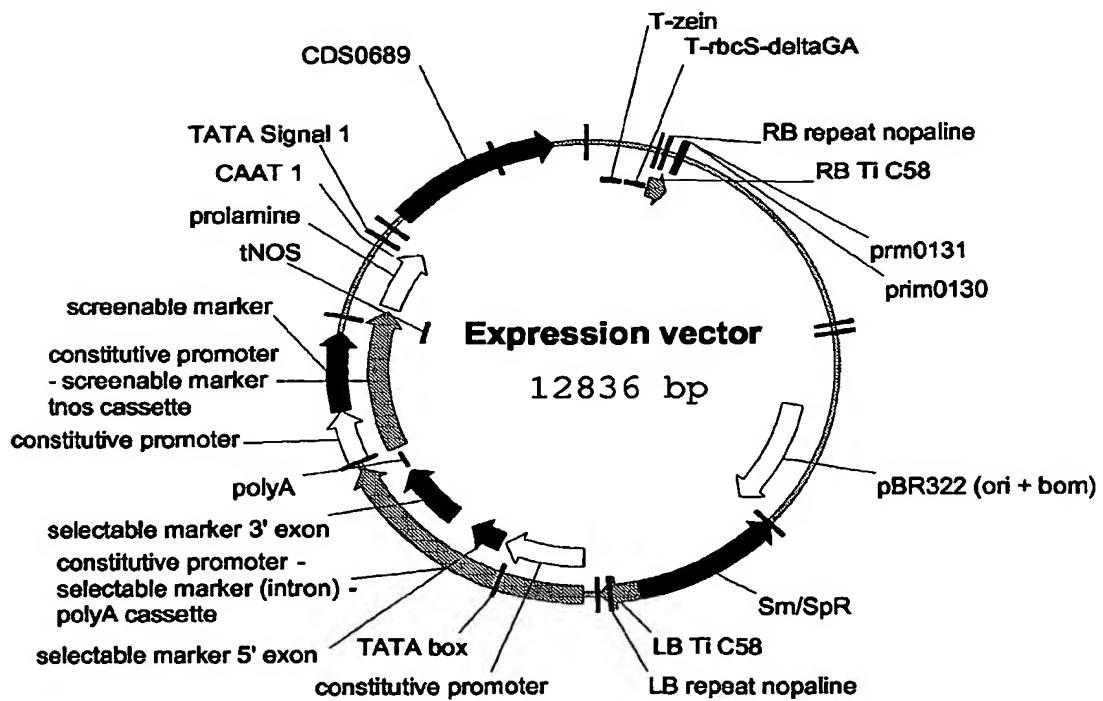
FIGURE 3

3/10

```
                                   401                                   440
                    CDS0689  (339) IQSSVVRKRSLPENDKDESKRNDKKRSLSVGKTRVSQTES
                 CDS0689 At  (294) TGEKDVRKRSLPEDEEKENHKRSEKR------RASDESN
   CDS0689 Medicago trunculata (284) GGD--ARKRSFSEN----------------------N
                 CDS0689 Os  (297) AAAAATAKRMAGSSKMRVIPSRYSITPGESLESSGAQERR
           CDS0689 Hv contig 123 (111) ISTASTCRRPAGSSKVRVVPSRYSMPGASLG-AATQDGR
            CDS0689 Ta contig  (1) -----------------GRYSLMPGASLGAASQERRR
           CDS0689 Zm partial  (87) TSNVATTKRPAGSSKVRVVPSRYSIPPGESLAAVTQGNRC
            CDS0689 Sacc sp 3'  (84) TSNAATAKRPAGSSKVRVVPSRYSITPGSYLAAVSQDKRS
        CDS0689 Pinus taeda 3'  (1) ----------------------------------
                  Consensus  (401)        KR      R      RYSL PGA LG
                                   441                                   480
                    CDS0689  (379) KNL---GEESRVKKRWEIPSEIVVHGNTESEKSPLSIIVK
                 CDS0689 At  (327) K------SEGRVKKRWEIPSEVDLYSSGEN--GDESPIVK
   CDS0689 Medicago trunculata (297) KGL---GSEIRAKKRWEIPIEEVDVSG--------FVM--
                 CDS0689 Os  (337) RKQSLPGSSGDANQNEEIRAKVISPSN--DPLSPQTISKV
           CDS0689 Hv contig 123 (150) RKESLPGSTGSTGQKEEIKAVPTDPVD--DDLSPESLDKV
            CDS0689 Ta contig  (21) KESLPGSIGGAGQKEEEIKAMPTSPVD--DDLSPESLDKV
           CDS0689 Zm partial  (127) KQS-----LPGSATETRVNLTEPPNDE----LSPEIIAKV
            CDS0689 Sacc sp 3'  (124) KQS-LPGPASAASQREEIRAKLTSPSK--DELSPETVAKV
        CDS0689 Pinus taeda 3'  (1) -----------XEARIVFGTGNSAIMAGGTKAPDTLERH
                  Consensus  (441) K     GS      R EI A   E      LSPESL KV
```

Motif 4

```
                                   481                                   520
                    CDS0689  (416) PDLPRIRIARCVNEILRDSSPAFRMIELIGKRSFFSSDE
                 CDS0689 At  (359) --ETPKIRTIRRVGGSPRDSGAAKRVAELQAK--------
   CDS0689 Medicago trunculata (324) ---LPKISTMRFVDESPRDSGAVKRVAELNGKRSVFCDED
                 CDS0689 Os  (375) AEMLPKIRTMPPPDESPRDSGCAKRVAELVGKRSFFTAAI
           CDS0689 Hv contig 123 (188) AELLPRIRTMPRPNETPRDSGCAKRAADLVGKRSFFAAAI
            CDS0689 Ta contig  (59) AELLPRIRTMPPPDETPRDSGCAKRAADLVGKRSFFAAAI
           CDS0689 Zm partial  (158) AELLPRIRTMPPSDESPRDSGCAKRVADLVGKRSFFTAAG
            CDS0689 Sacc sp 3'  (161) AELLPRIKTMPASDESPRDSGCAKRVADLVGKRSFFTXAI
        CDS0689 Pinus taeda 3'  (29) KMKIPKIKTVRFTTESPRDSGCIKREIDRIGKRSFFAPDG
                  Consensus  (481) AELLPRIRTMP  DESPRDSGCAKRVADLVGKRSFF AAA
                                   521                                   560
                    CDS0689  (456) --------DKEPPVCQVLSFAEEDAEEE-----------
                 CDS0689 At  (389) --------DRNFTFCQILKFEE----------------
   CDS0689 Medicago trunculata (361) EEERVMVEEEGGSVCQVLNFAEDDDDDDDYGEQG------
                 CDS0689 Os  (415) ------------EDGRALDVEAPEAVAEA----------
           CDS0689 Hv contig 123 (228) AGD---GSAISSYQARVLEAFAPEEAAAAGALSDEAAAAG
            CDS0689 Ta contig  (99) AGD---CSAISSYQARVLEAFAPEEAAAAAEALGDEAASA
           CDS0689 Zm partial  (198) DDG----NLVTPYQARVVELESPEAAAEEAEA--------
            CDS0689 Sacc sp 3'  (201) EDG----NFVTPYQAPVGEL------------------
        CDS0689 Pinus taeda 3'  (69) ITST--PSIDXXDAGKPLRRESVHEIXHAXXX--------
                  Consensus  (521)            ARVLE EA E
                                   561                         592
                    CDS0689  (476) ------------------------------
                 CDS0689 At  (403) ------------------------------
   CDS0689 Medicago trunculata (395) ------------------------------
                 CDS0689 Os  (432) ------------------------------
           CDS0689 Hv contig 123 (265) ALSDEAAAAAAAEALSDEAAAAEALSDEAAA
            CDS0689 Ta contig  (136) GEALGDEAAA--------------------
           CDS0689 Zm partial  (226) ------------------------------
            CDS0689 Sacc sp 3'  (217) ------------------------------
        CDS0689 Pinus taeda 3'  (99) ------------------------------
                  Consensus  (561)
```

FIGURE 3 (continued)

SEQ ID NO 1: *Nicotiana tabacum* seedy1 coding sequence (CDS0689)

atgagtgtgttacaatacccagaagggattgacccagcagatgttcagatatggaacaatgc
agcatttgataatggagattctgaagatttgtcttcgctgaaacgttcttggtctcctctga
aacccctttcggttaggccatcagattcctttgaatctgatttgtcaagtaaggaaaatcaa
actcctttatttgagaattcatctgttaatctctcatctccgttacccataaagccacttaa
ccctaatggggctctggaaaattcaagactcaagccgaacaagcccaattccaaacagagtc
ttgatgagatggcggctagaaagagcggaaagggaaatgatttccgtgatgagaagaaaata
gacgaggaaattgaagaaattcagatggagattagtaggttgagttcaagattagaggcttt
gagaattgaaaaggctgagaaactgttgctaagactgttgaaaagcgaggaagggttgtgg
cagcaaagtttatggagccaaaacaaagtgttattaagattgaagagcgtatatcaatgagt
gcaagaacaaaggtggagcagagaaggggtcttagtttaggaccatctgagatttttactgg
aacgcggcggcgagggttgagtatggggccatcagatattctagcagggacaacaaaggcac
ggcaattgggaaagcaagagatgattattactcctattcagccaatacaaaacaggcgaaag
tcgtgttttggaagcttcaagagattgaagaagagggaaaaagttcaagccttagtcctaa
atcaagaaaaactgctgcaagaacaatggttacaacaaggcaggcagttactacaattgcat
caaagaagaatttgaaaaaagatgatggactttttgagttcagttcagccaaagaagttgttt
aaagatctcgaaaagtctgctgctgctaataagaagccccagaggccggggagggttgtggc
tagtaggtataatcagagtacaattcagtcatcagtagtgagaagaggtctttacctgaaa
atgataaggatgagagtaagagaaatgataagaaacggtcgttatctgtagggaaaacgcgt ·
gtgtctcaaactgagagcaagaatttgggtactgaaagtagggtgaaaaagagatgggaaat
tcctagtgagattgtagttcatggaaacacagagagtgagaaatctccactaagcattattg
tgaagcctgatttgcttccgcgaattaggattgctcggtgtgtgaatgagactcttagggat
tctggacctgctaaaagaatgatagagttgataggcaagaaatcgttttttcagtagtgatga
agataaggagccacctgtctgtcaagtttttaagttttgcagaggaagatgctgaagaggaat
aa

SEQ ID NO 2: *Nicotiana tabacum* seedy1 protein (CDS0689)

MSVLQYPEGIDPADVQIWNNAAFDNGDSEDLSSLKRSWSPLKPLSVRPSDSFESDLSSKENQ
TPLFENSSVNLSSPLPIKPLNPNGALENSRLKPNKPNSKQSLDEMAARKSGKGNDFRDEKKI
DEEIEEIQMEISRLSSRLEALRIEKAEKTVAKTVEKRGRVVAAKFMEPKQSVIKIEERISMS
ARTKVEQRRGLSLGPSEIFTGTRRRGLSMGPSDILAGTTKARQLGKQEMIITPIQPIQNRRK
SCFWKLQEIEEEGKSSSLSPKSRKTAARTMVTTRQAVTTIASKKNLKKDDGLLSSVQPKKLF
KDLEKSAAANKKPQRPGRVVASRYNQSTIQSSVVRKRSLPENDKDESKRNDKKRSLSVGKTR
VSQTESKNLGTESRVKKRWEIPSEIVVHGNTESEKSPLSIIVKPDLLPRIRIARCVNETLRD
SGPAKRMIELIGKKSFFSSDEDKEPPVCQVLSFAEEDAEEE

FIGURE 4

SEQ ID NO 3: *Oryza sativa* seedy1 coding sequence

atggaggaggacccgctcatcccgctggtccacgtctggaacaacgccgccttcgacgactc
ctcgtgttccagatcggcttggctcccccaaagccccgccgtcgcggccgtccgcaagggcg
acaaggagaatcaccgccccgaggttgttgatgtcgccgccggctacgacgtcgaggccgag
atcggccacatcgaggcggagatcctgcgcctctcgtcccggctccaccatctccgcgtctc
caagcagccggagcccaaccgcgacgacgctccgatggggggagatggtcgcgaaggtgaggc
cccggccgaggggcctcagcctcgggcccctggatgtgatctccatcgtcaatcgtgagaag
catccgctgcgcaccaagcagcctccggcgacgcggggcaggggctcagcctcgggcccat
ggagatcgccgcggcgaaccctagggtgcccgcggcggcgcagcatcagcaacagcaacgcg
ctggcacggcgcggatcctgaagccaatcaaggagcctccggtgcagcgtcgcagggggcgtc
agcctcgggccgttggagatccaccacggcgtcggcagcaaggcaccagcggcggcgcgagc
caagccgttcaccaccaagctcaacgccattcgagaagaaacccgaccctccaagcaattcg
ccgtccccgccaagccatggccgtcgagcaatacaaggcagacactggactcgaggcaagga
acagcagcaagtcgagcgaaggcgaggagcccgagcccaggcccaggaggcaatccaatgg
caaggctactgacacaaggggaggcaacaaggtggtggatgagctcaagcccaaaggtgcgt
cgtcaagtcagagcggcagcgccgccgccgccgccactgccaagaggatggcggggagctcc
aagatgagggtcatcccgagccgctacagcctcactcctggcgcttcccttggaagcagtgg
agcacaggagaggcgacgcaagcagtctctcccaggatcatcaggggatgcgaaccagaatg
aggaaatcagagcgaaggtcatcgagccttccaatgatccactctctcctcaaacgatctcc
aaggttgctgaaatgctcccaaagatcaggaccatgccgcctcctgacgagagccctcgcga
ttccggatgcgccaagcgggttgccgaattggtcgggaagcgctcgttcttcacggctgcag
ccgaggacgggcgggcgctcgacgtcgaagcacccgaggcggtcgcagaagcttgagatgaa
ccaccatggtttgatccgttccttccatcagctc

SEQ ID NO 4: *Oryza sativa* seedy1 protein

MEEDPLIPLVHVWNNAAFDDSSCSRSAWLPQSPAVAAVRKGDKENHRPEVVDVAAGYDVEAE
IGHIEAEILRLSSRLHHLRVSKQPEPNRDDAPMGEMVAKVRPRPRGLSLGPLDVISIVNREK
HPLRTKQPPATRGRGLSLGPMEIAAANPRVPAAAQHQQQQRAGTARILKPIKEPPVQRRRGV
SLGPLEIHHGVGSKAPAAARAKPFTTKLNAIREETRPSKQFAVPAKPWPSSNTRQTLDSRQG
TAASRAKARSPSPRPRRQSNGKATDTRGGNKVVDELKPKGASSSQSGSAAAAATAKRMAGSS
KMRVIPSRYSLTPGASLGSSGAQERRRKQSLPGSSGDANQNEEIRAKVIEPSNDPLSPQTIS
KVAEMLPKIRTMPPPDESPRDSGCAKRVAELVGKRSFFTAAAEDGRALDVEAPEAVAEA

FIGURE 4 (continued)

SEQ ID NO 5: *Medicago trunculata* **seedy1 encoding sequence**
```
aaaaacgttaaggactaaaaatataataaaatttaagtagggattcataatggaagcacccc
tatttacagggatcttaaatataattaaccctaatatttatgacagaaacccttttgaaatc
acatcggagcgtgtatgagtagccgtttcacatccaacggccagtaagagcgtaactttatt
tcttccctcttcaatctccaacggtcacataatctcttccaaatacaaataattccctcttt
caacctcactcttcatttcttcaacccaaacccaaaaaactaatcagattcttcttaaatct
tgaaacctttctcccaaaagcacttaaataaaaagcacttaaccatgaataacacaaacaa
caacaacattcttcttcattccacacaggttcaagtgtggaacaacgcagcattcgatggtg
aagatttcgccatgaattcatcttctgattccatcaaagagaatctaaacccatccgcattc
aacattgttccttcttcaaacaaaagaactattgatgatgaaattgcggaaattgaaagtga
aattaagcgattaacttcgaagctggaattgcttcgtgttgaaaaagctgaagaaaaatcg
cttctgaaaagcgtgttagtggaattggtactggaagaatagtagcagcgaagtttatggaa
ccgaagaaaaacgttacaccgaaacgaaacggtgtcgttttcaaggaggagacaccgaaacg
aaacggtgtcgtttcggatacgccgaaatctagggttaattggagaagagggatgagtttag
gtccgatggagattgccgggaaagtgatggcaccgccggcgatgacgattactccggcgacg
gtgaatcggaggaagtcttgtttctggaaaccgcaggaaagttgtgaagtaatgccgtcggg
gattactccggcgacggtgaataggaggaaatcttgttttttgaaacctcaagaaagttgtg
aagaaatcgaagaaaaacgatttgcaaaccgaatttgaatttgaattcaaattcagttaat
tctgcggttggatcgattaagcgtgtgaagaagaaagatgaagaaattgctcaggttcaacc
gaagaagctgtttgaaggtgaaaaatcagtgaagaaatcgttgaaacaaggtagaattgttg
caagccggtataattccggtggtggtggtggtggtgatgcgaggaaaagatcgttttcggagaat
aataagggtttagggagtgaaatcagggctaagaagagatgggagataccaattgaagaagt
ggatgtgagtggtttttgttatgttaccgaagatttcgacaatgaggtttgttgatgagagtc
ctagagattctggtgctgttaaaagagttgctgaattgaatggaaaaagatcttacttttgt
gatgaagatgaggaggagagagtgatggtggaggaagaaggtggttctgtttgtcaggtttt
gaattttgctgaagatgatgatgatgatgatgattatggtgaacaagggtaattgtggaaat
tggaattgatttgttttttgtggggttgtgtggaactggctatgttctgcttgattcttttgc
attttggtgtgaaactaaagatgaggtgaaaagtttatgcttgttaaattggattggtttat
atgttttgaaataataacaacaagcatgtgtcttgcttaataattgtatattgtttttgtttg
ttttataatgatatggatttaatttgtatacacaatataatatagtatgcattgagagagtt
tttcgttcagtattcattctgattttagtgtttatctcattctagaagattgtattttgttg
```

SEQ ID NO 6: *Medicago trunculata* **seedy1 protein**
```
MNNTNNNNILLHSTQVQVWNNAAFDGEDFAMNSSSDSIKENLNPSAFNIVPSSNKRTIDDEI
AEIESEIKRLTSKLELLRVEKAERKIASEKRVSGIGTGRIVAAKFMEPKKNVTPKRNGVVFK
EETPKRNGVVSDTPKSRVNWRRGMSLGPMEIAGKVMAPPAMTITPATVNRRKSCFWKPQESC
EVMPSGITPATVNRRKSCFLKPQESCEENRRKTICKPNLNLNSNSVNSAVGSIKRVKKKDEE
IAQVQPKKLFEGEKSVKKSLKQGRIVASRYNSGGGGGDARKRSFSENNKGLGSEIRAKKRWE
IPIEEVDVSGFVMLPKISTMRFVDESPRDSGAVKRVAELNGKRSYFCDEDEEERVMVEEEGG
SVCQVLNFAEDDDDDDDYGEQG
```

FIGURE 4 (continued)

SEQ ID NO 7: *Saccharum sp.* **encoding sequence (partial 5' end)**
Cgcaccgcgagtttcgaaaaaccaacctatcgcgcctcagatcacgcgaggacgcgagggga
agcaggaatccctccgctcccagccgcctcctccgctcacccatcgatcgatcgtccgtccg
gtccaggggggctctccggcggcggtggcgatggaggaggacccgctcatcccgctggtgcac
gtctggaacaacgccgccttcgaccacgcctcctcctccgcgtggcacgcccactcccctgt
gcccgcgagcgcacgtcgcgaggcggagggggacaaggagaaccaccgccccgaccccgacc
ccgacgtcgaggcggagatcggccacatcgaggcggagatcctgcgcctgtnctcccgcctn
caccaccttcgcacctccaagcagtcggagccgtccaagcgcggagaggtcgcgcccgcgcc
cgcggcgaaggcgaaagcggcggcggcggcgcggctgcggacgcggggggctcagcctgggcc
cgctcgacgtcgccgctgccggtaaccccaacccgctcaccaccgacaaccagcagcagcag
ccgcgtgccgcgcagggtctgaagccgatcaagcaggccacggcggcggcgggcaagggcgt
aagacttgggcccccttcgacatggtcggcgcgaaccctagggtccctccgcccn

SEQ ID NO 8: *Saccharum sp.* **seedy1 protein (partial N-term)**
MEEDPLIPLVHVWNNAAFDHASSSAWHAHSPVPASARREAEGDKENHRPDPDPDVEAEIGHI
EAEILRLXSRLHHLRTSKQSEPSKRGEVAPAPAAKAKAAAAARLRTRGLSLGPLDVAAAGNP
NPLTTDNQQQQPRAAQGLKPIKQATAAAGKGVRLGPLRHGRR

SEQ ID NO 9: *Zea Mays* **seedy1 encoding sequence (partial 3' end)**
ccacgcgtccggccgttcgagaggaggaaggccagcgttccaaggagcacgccgtccccgcc
agaccgtggccatccagcaatgccaggcacccactggatgccaggcaaggcaccgcagcaag
cagagccaaggcgaggagcgggagcataagccccagcaggttcaggaggcagtccacttcca
aggctgccgagacaagagcgggaaatgccaagcctacagaggcgacgaggggagggagcgaa
gcggtcaatcacaccagcaatgtagccacgacgaagaggccggcggggagctccaaggtcag
ggttgtcccgagccgctacagcatcccacctggctcctccctagcagctgtgacacaaggca
accgatgcaagcagtctctcccaggatcggctactgagaccagagtaaatctcactgagccg
ccgaacgacgagttgtctcctgaagaacttgccaaggttgcagagctgctcccaaggattag
gaccatgccgccttctgatgagagcccgcgtgactcgggatgtgccaagcgtgttgctgatt
tggtcgggaagcgatccttcttcactgctgcaggggacgatggcaatctcgttacgccctac
caggcacgggtggttgaacttgaatcacccgaggcagcagcagaagaagcagaagcttgaga
agtttgtctttgatcaattccgaagtggcttgcatctgggcgtggcctctttttgcagtgtg
tgctactacatagtctactgttacattcatatcatatcacatttcctattttttcccccttg
agacattgcttagtacttttgtgttgccttgtgaaaagagagtggaaggttcatctgctgat
nccttgtt

SEQ ID NO 10: *Zea Mays* **seedy1 protein (partial C-term)**
TRPAVREEEGQRSKEHAVPARPWPSSNARHPLDARQGTAASRAKARSGSISPSRFRRQSTSK
AAETRAGNAKPTEATRGGSEAVNHTSNVATTKRPAGSSKVRVVPSRYSIPPGSSLAAVTQGN
RCKQSLPGSATETRVNLTEPPNDELSPEELAKVAELLPRIRTMPPSDESPRDSGCAKRVADL
VGKRSFFTAAGDDGNLVTPYQARVVELESPEAAAEEAE

FIGURE 4 (continued)

SEQ ID NO 11: *Arabidospis thaliana* **seedy1 coding sequence**
atgacatcaattgaggcaacagaaacgcttaacgctcctccaaagcttcagatctggaacaa
cgctgccttcgacgatggagattctcaaatcacttccgccatcgaagcttcttcttggtctc
acctcaacgaatcattcgattccgattgtagcaaggagaatcagtttccgatttcggtttcc
tcttcgctccaatcctcagtctcgatcaccgaagctccgtcagcaaaatccaagaccgtgaa
gaccaaatccgccgcagatcggagtaaaaagcgagatatcgatgcagagatcgaagaagtag
agaaggagatcggacgattatcgacgaaattggagtcgctccgattagagaaggcggagcaa
accgcaagaagcattgctatacgtggaagaatcgttccggcgaagttcatggaatcatctca
gaaacaagtgaaattcgacgattcgtgttttacaggatcgaaatcaagagccactcgtagag
gcgttagtcttggaccagcggagatattcaattccgcgaagaaatctgaaactgtgactcct
cttcaatcagctcagaatcgacgcaagtcttgtttctttaagcttcctggaatcgaagaagg
tcaagtgacgacacgaggtaaaggaagaacgagtttgagtctgagtccgagatctcgcaaag
cgaaaatgacggcagctcagaagcaagcagctacgacggtggggtcaaagagagctgtgaag
aaagaagaaggagttctcttaacaatccagcctaagaggctattcaaagaagatgaaaagaa
tgtttctttaaggaaaccattgaaaccaggaagagttgtggctagtaggtacagtcaaatgg
gtaaaacgcagactggagagaaagatgttaggaaaaggtcgttgcctgaggatgaagagaaa
gagaatcataagaggtcggagaagagaagagcttctgatgaaagtaacaagagtgaagggag
agtgaagaagagatgggagattccaagtgaagttgatctgtatagcagtggtgagaacggtg
acgagtctcctatagttaaggagctacctaagatcagaacgcttcgtcgtgtgggagggagc
cctcgtgattcaggtgctgctaagagagttgcagaattacaagccaaggatcgtaacttcac
tttttgccagcttctgaagtttgaagaatgaatgatccgcttatcaatttgagtaaaatcca
caactcttgttgtggtt

SEQ ID NO 12: *Arabidospis thaliana* **seedy1 protein**
MTSIEATETLNAPPKLQIWNNAAFDDGDSQITSAIEASSWSHLNESFDSDCSKENQFPISVS
SSLQSSVSITEAPSAKSKTVKTKSAADRSKKRDIDAEIEEVEKEIGRLSTKLESLRLEKAEQ
TARSIAIRGRIVPAKFMESSQKQVKFDDSCFTGSKSRATRRGVSLGPAEIFNSAKKSETVTP
LQSAQNRRKSCFFKLPGIEEGQVTTRGKGRTSLSLSPRSRKAKMTAAQKQAATTVGSKRAVK
KEEGVLLTIQPKRLFKEDEKNVSLRKPLKPGRVVASRYSQMGKTQTGEKDVRKRSLPEDEEK
ENHKRSEKRRASDESNKSEGRVKKRWEIPSEVDLYSSGENGDESPIVKELPKIRTLRRVGGS
PRDSGAAKRVAELQAKDRNFTFCQLLKFEE

FIGURE 4 (continued)

**SEQ ID NO 13: Sequence of the [PRO0090 – CDS0689 – terminator]
expression cassette**

cttctacatcggcttaggtgtagcaacacgactttattattattattattattattatt
attttacaaaaatataaaatagatcagtccctcaccacaagtagagcaagttggtgagttat
tgtaaagttctacaaagctaatttaaaagttattgcattaacttatttcatattacaaacaa
gagtgtcaatggaacaatgaaaaccatatgacatactataattttgttttttattattgaaat
tatataattcaaagagaataaatccacatagccgtaaagttctacatgtggtgcattaccaa
aatatatatagcttacaaaacatgacaagcttagtttgaaaaattgcaatccttatcacatt
gacacataaagtgagtgatgagtcataatattattttctttgctacccatcatgtatatatg
atagccacaaagttactttgatgatgatatcaaagaacatttttaggtgcacctaacagaat
atccaaataatatgactcacttagatcataatagagcatcaagtaaaactaacactctaaag
caaccgatgggaaagcatctataaatagacaagcacaatgaaaatcctcatcatccttcacc
acaattcaaatattatagttgaagcatagtagtaatttaaatcaactagggatatcacaagt
ttgtacaaaaaagcaggctggtaccggtccggaattcccgggatatcgtcgacccacgcgtc
cgctgacgcgtgggttccactacatcaagacatctactacactcatcttttttgcacttatt
gggtgtaaattttttgaaacccagttgagaaaaatgagtgtgttacaatacccagaagggatt
gacccagcagatgttcagatatggaacaatgcagcatttgataatggagattctgaagatttt
gtcttcgctgaaacgttcttggtctcctctgaaacccctttcggttaggccatcagattcct
ttgaatctgatttgtcaagtaaggaaatcaaactcctttatttgagaattcatctgttaat
ctctcatctccgttacccataaagccacttaaccctaatggggctctggaaaattcaagact
caagccgaacaagcccaattccaaacagagtcttgatgagatggcggctagaaagagcggaa
agggaaatgatttccgtgatgagaagaaaatagacgaggaaattgaagaaattcagatggag
attagtaggttgagttcaagattagaggctttgagaattgaaaaggctgagaaaactgttgc
taagactgttgaaaagcgaggaagggttgtggcagcaaagtttatggagccaaaacaaagtg
ttattaagattgaagagcgtatatcaatgagtgcaagaacaaaggtggagcagagaaggggt
cttagtttaggaccatctgagatttttactggaacgcggcggcgagggttgagtatggggcc
atcagatattctagcagggacaacaaaggcacggcaattgggaaagcaagagatgattatta
ctcctattcagccaatacaaaacaggcgaaagtcgtgtttttggaagcttcaagagattgaa
gaagagggaaaaagttcaagccttagtcctaaatcaagaaaaactgctgcaagaacaatggt
tacaacaaggcaggcagttactacaattgcatcaaagaagaatttgaaaaaagatgatggac
ttttgagttcagttcagccaaagaagttgtttaaagatctcgaaaagtctgctgctgctaat
aagaagccccagaggccgggggagggttgtggctagtaggtataatcagagtacaattcagtc
atcagtagtgagaaagaggtctttacctgaaaatgataaggatgagagtaagagaaatgata
agaaacggtcgttatctgtagggaaaacgcgtgtgtctcaaactgagagcaagaatttgggt
actgaaagtagggtgaaaaagagatgggaaattcctagtgagattgtagttcatggaaacac
agagagtgagaaatctccactaagcattattgtgaagcctgatttgcttccgcgaattagga
ttgctcggtgtgtgaatgagactcttagggattctggacctgctaaaagaatgatagagttg
ataggcaagaaatcgttttttcagtagtgatgaagataaggagccacctgtctgtcaagtttt
aagttttgcagaggaagatgctgaagaggaataatgtgtaataaagggagctgctaactctt
ttcatgctctttcaattttcaatcctgccttttaattttttgttcattcgtgccttttaattg
aatggggaagcattctttttgcttcctcaaactggtattctagcttctgaattacattgtatg
gtacaatatgaataaggttttgtcttccggcaggttgtccaagttagttttttagcttaaaat
agatgcggcagcggccgctctagagtatccctcgaggggcccaagcttacgcgtacccagct

FIGURE 4 (continued)

ttcttgtacaaagtggtgatatcacaagcccgggcggtcttctagggataacagggtaatta
tatccctctagatcacaagcccgggcggtcttctacgatgattgagtaataatgtgtcacgc
atcaccatgggtggcagtgtcagtgtgagcaatgacctgaatgaacaattgaaatgaaaaga
aaaaaagtactccatctgttccaaattaaaattcattttaaccttttaataggtttatacaa
taattgatatatgttttctgtatatgtctaatttgttatcatccgggcggtcttctagggat
aacagggtaattatatccctctagacaacacacaacaaataagagaaaaaacaaataatatt
aatttgagaatgaacaaaaggaccatatcattcattaactcttctccatccatttccatttc
acagttcgatagcgaaaaccgaataaaaaacacagtaaattacaagcacaacaaatggtaca
agaaaaacagttttcccaatgccataatactcgaac

SEQ ID NO 14: rice prolamin RP6 promoter sequence
ccttctacatcggcttaggtgtagcaacacgactttattattattattattattattattat
tattttacaaaaatataaaatagatcagtccctcaccacaagtagagcaagttggtgagtta
ttgtaaagttctacaaagctaatttaaaagttattgcattaacttatttcatattacaaaca
agagtgtcaatggaacaatgaaaaccatatgacatactataattttgttttttattattgaaa
ttatataattcaaagagaataaatccacatagccgtaaagttctacatgtggtgcattacca
aaatatatatagcttacaaaacatgacaagcttagtttgaaaaattgcaatccttatcacat
tgacacataaagtgagtgatgagtcataatattattttttcttgctacccatcatgtatatat
gatagccacaaagttactttgatgatgatatcaaagaacatttttaggtgcacctaacagaa
tatccaaataatatgactcacttagatcataatagagcatcaagtaaaactaacactctaaa
gcaaccgatgggaaagcatctataaatagacaagcacaatgaaaatcctcatcatccttcac
cacaattcaaatattatagttgaagcatagtagtagaatccaacaaca

SEQ ID NO 15: Motif 1 CORE SEQUENCE
WXNAXXD

SEQ ID NO 16: Motif 2 CORE SEQUENCE
KENXXP

SEQ ID NO 17: Motif 3 (coiled coil) CORE SEQUENCE
$EX_{1-6}EXXRLXXXLXXLR$

SEQ ID NO 18: Motif 4 CORE SEQUENCE
$LPXIX_{1-10}RDSGXXKRX_{1-6}K$

FIGURE 4 (continued)